DOCUMENT RESUME

ED 066 933 FL 002 356

AUTHOR Lee, Richard R.
TITLE The Social Evaluation of Speech: Implications from the Laboratory to the Classroom.
PUB DATE 5 Mar 71
NOTE 10p.; Paper presented at the Fifth Annual TESOL Convention, March 5, 1971, New Orleans, La.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Communication (Thought Transfer); *Communication Skills; Connected Discourse; Experiments; Language Instruction; Language Programs; *Lower Class; *Nonstandard Dialects; Racial Factors; Role Perception; Social Factors; *Social Mobility; *Speech; Speech Education; Standard Spoken Usage; Tables (Data)

ABSTRACT

The teaching of dialect modification has been justified by the barrier postulate which says that lower class speech prevents upward social mobility. When translated into a testable hypothesis, the barrier assumption has not been confirmed. The movement called dialect modification did not arise from empirical research in social cognition, but instead from the intuitions of scholars trained in linguistics. What is different in the communication patterns of the poor goes beyond the phonology and syntax of the sentence and linguistic analysis to the broad basis of communication skills and points out the need for communication training. (Author/VM)

FILMED FROM BEST AVAILABLE COPY

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

# THE SOCIAL EVALUATION OF SPEECH

## IMPLICATIONS FROM THE LABORATORY TO THE CLASSROOM

Richard R. Lee

ED 066933

The teaching of dialect modification has been justified by the barrier postulate, which says that lower class speech prevents upward social mobility. When translated into a testable hypothesis the barrier assumption has not been confirmed. In fact, neither the variables of race, nor social class accurately predict the evaluation of speech samples by listeners. The barrier postulate has not been proven right, and epistemologically, it cannot be proven wrong. For this and other reasons, the teaching of dialect modification should be redirected to the teaching of communication competence in specific situations.

One of the working assumptions in TESOL is that the speech of lower class speakers is a barrier to social mobility. In his presidential address in 1970, David P. Harris noted the shifting attention of the organization and a growing concern "for those many thousands of American children and adults whose academic success and social mobility are severely restricted by the kind of English they use."[1] This concern has been translated into a special kind of instruction for the children of the poor, instruction called alternatively dialect modification or standard English as a second dialect. The postulate which justifies such instruction has never been adequately tested, that is, that the speech of members of the lower classes constitutes a barrier. Implicit in Harris' observation are several assumptions about the relationship between speech and social perception which must be made explicit before the barrier postulate can be tested.

FL 002 356

The first assumption is that differences exist in the speech of different social classes in America, differences which are best described as differences in dialect, specifically differences in phonology and syntax. One compilation of these phonological and syntactic differentia has been made by Raven McDavid.[2] It should be noted in passing that there are many other ways to describe speech differences: sentence length, word choice, type-token ratios, appropriateness of responses to an interview question, and so on.

A second assumption is that untrained listeners detect and isolate these phonological and syntactic differentia within a message, a detection which negatively influences the listener's evaluation of the social worth of the speaker. Stated another way, this assumption says that how a speaker says something is perceived independently of what he says, and has priority in the process of forming social judgments. A third assumption is that these differentia provide for the casual listener a reliable cue to a speaker's social class since, it is further assumed, that these differentia are common to virtually every member of a particular social class or ethnic minority. In other words, thirty listeners acting independently will all come to about the same conclusion about a speaker's rank in the social heirarchy. Inherent in the barrier postulate is yet another assumption, that for the average listener, the identification of the speaker's social class is a primary percept, a cognition formed early in an encounter. These assumptions constitute the barrier postulate and provide a rationale for instruction in dialect modification.

There is a good reason to believe that speech differences, however categorized, have little influence in the formation of impressions about strangers, particularly when the influence of formal features of speech are weighed against the influence of content. Victor Cline, a psychologist who works in person perception, presented information about interviewees in a variety of ways to several panels of judges. He showed them silent and sound motion pictures of the interviews: he played just the sound track alone: some judges read a manuscript of the interview and others heard very short samples of speech that were essentially content free. Cline concluded that "what most interviewees say in response to the interview questions is far more important, as cues, than what they look like, what the voice sounds like [and] how they act or move (without the sound) all put together.[3] Cline's findings are strikingly different from those reported by Putnam and O'Hearn[4], by Labov[5], by Harms[6], and by Tucker and Lambert[7], all of whom have reported finding a strong association between a speaker's real social

class and listeners' judgment about their social class after listening to their speech. This difference can probably be ascribed to different experimental techniques. Cline did not control for content. He thus avoided creating the artificial listening situation of repeated, identical messages. Anyone who has stood in a reception line can testify how stunningly monotonous it is to hear the same message over and over again. By using real interviews with their inherent variability of content, Cline's elicitation technique did not eliminate potentially powerful independent variables. Finally, Cline did not focus his listeners' attention on his independent variables: his listeners were asked to predict how each interviewee behaved in real life and what kind of personality he had. Cline, however, did not sample from the lower socio-economic strata. It is against this part of society that the barrier supposedly stands. The barrier postulate can be translated into a prediction about listener behavior. It says, first of all, that a group of listeners will generally agree with one another in their ratings of several speakers. Individual data from any experiment in social judgment must pass this hurdle first before pooled data can be said to have any significane. Without clearing the hurdle of inter-rater reliability, pooled data, expressed as mean scores, can possibly characterize behavior of the entire group but of no individual within it.[8] Since social cognition is a private act performed without consultation with the group and since the barrier postulate also suggests that social perception is a stable process, judgments about a set of voices should be essentially identical from one week to the next. More precisely, speech stimulated social cognition should demonstrate test-retest reliability.

The barrier postulate also suggests an appropriate dependent measure. Since social mobility is partially a function of job availability, listeners should be asked to make judgments about the highest job a speaker could hold. Martin Joos attributes keen social perception to very young children. "Long before any teacher began to correct his English, the child has learned

all he needs to know, at his age, about people and their places: he has developed considerable skill in judging adults by their speech."[9] The barrier postulate suggests an appropriate dependent measure and Joos' idea suggests where to begin.

Twenty very average second graders were asked to rank eight occupations on the basis of the speech qualification demanded by each one. Early in the development of the instrument, we were concerned that children might confuse speech qualification with occupational status. We needn't have worried. Given eight occupations, these seven-year-olds regularly ranked the occupation of artist last, below waitress and truckdriver. These youngsters showed clear consensus that a television announcer required better speech than a teacher, that teaching required better speech than working in an office, and that a waitress and a truckdrive required the least qualification of all.

The consensus evaporated when they were asked to rank each voice according to the highest job that the speaker could hold. On the retest three weeks later, individual ratings of a voice changed, either up or down, in almost every case. Inter-rater reliability within the class was -.03. These results suggest that the rating task for these second graders was merely guessing behavior, even though the tests were administered individually by a woman with long experience with young children. If there is any truth to Joos' notions about the acuity of children's ability to judge speech, it is not to be found here.

The speech samples themselves were elicited from interviews conducted with eight subjects, between 15 and 17 years old, four boys and four girls, four blacks and four whites, and four from families making more than $5,000 a year and four speakers from families earning less than $3,000 a year. To control regional variations in the samples, all subjects were born and reared within 25 miles of the same small southern city. Notice that the

independent variables are race, sex and socio-economic class, not dialect. This seems to be the most straightforward interpretation of the barrier postulate and avoids begging the question whether in fact dialect is an independent variable. Each speech sample contained two or three questions by an adult interviewer and the subjects' replies, which were edited to remove long pauses between question and reply or within the reply. Disfluencies such as "uhm" and bad strats were also removed. The eight edited samples were from 35 to 45 seconds long. The purpose was to assemble a sample of fairly continuous speech, not to present a representative sample of an entire interview. What was presented was a sample of speech, not a sample of communication competence.

After analyzing the results of the second graders, additional pilot work was done with seventh graders, ninth graders, and high school seniors. Inter-rater reliability rose somewhat, but to nothing approaching consensus.

Table I

Individual Inter-Rater Reliability, Public School Listeners, Edited Speech Samples.

| | Test | Retest |
|---|---|---|
| 7th grade, racially mixed (N=19) | -.03 | .03 |
| 9th grade, black, (N=19) | .03 | .00 |
| 9th grade, white, (N=38) | .24 | |
| 11th grade, racially mixed (N=42) | .31 | |
| 12th grade, white (N=19) | .18 | .32 |

Instead of using the customary tests of statistical significance, an inter-rater reliability coefficient of .70 was determined to be the lowest acceptable value. A correlation attempts to account for variance among ratings. A coefficient of .70 in inter-rater reliability would account for slightly more than half of the variance which in turn would suggest that characteristics of the stimuli, whatever they might be, influenced ratings at least as much as all other factors combined, such as idiosyncratic characteristics of the raters. A

coefficient of this magnitude should not be hard to attain for judgments based on a four rank scale, or so it seemed at the time. It was decided to take the experiment to an older more educated group of listeners.

A group of fifty college sophmores, juniors, and seniors listened to the same eight voices in random order under three conditions: the edited samples, the same samples in unedited form, which were inevitably a few seconds longer, and the same voices again reading a short narrative passage in a third condition. In other words, the listeners rendered 24 ratings on eight voices according to the highest job that each speaker could hold on the basis of his speech. The listeners returned a week later and rated the same eight voices under the same three conditions again. The purpose behind having three conditions was to measure the influence of different elicitation conditions upon listener ratings. As it turned out, the data never cleared the hurdle of inter-rater reliability, so nothing conclusive can be said about these differences.

Two trends appeared in the data. Mean ratings for edited speech were somewhat higher than for either the unedited samples or the reading passages. Apparently pauses and hesitations have a slight negative effect on social cognition. Another trend appeared in a comparison between the test and retest rankings. The mean retest ratings for all three conditions were higher than the originals, although it would be pointless to speculate which setting resembles casual social cognition more, the test or the retest. Experience with the experimental task is an independent variable itself and must be taken into account when interpreting one-shot experiments.

Table II

Individual Inter-rater Reliability College Listeners (N=50)

| | Test | Retest |
|---|---|---|
| Edited Samples | .30 | .50 |
| Unedited samples | .52 | .58 |
| Reading passages | .48 | .56 |

More to the point, however, is a comparison of the ratings given by individuals to the eight voices during the first session and ratings they gave a week later. Under the same conditions, these college students awarded the same rating to the same voice a second time on an average of 4.65 times out of eight, or a little more than half the time, on a scale with only four ranks. Thus neither inter-rater nor intra-rater reliability could be established.

There is nothing in these data to support the barrier postulate, that is, that the speech of any social class is reliably evaluated in a particular fashion. By speech, I mean syntax, phonology, content, word choice and any other isolable formal feature of connected discourse. There was little concensus within any of the groups tested, and individual ratings fluctuated from one week to the next. Reasonable inter-rater reliability might be obtained by reducing the rating to a zero-one, good-bad judgment. Such a rating has no relevance to either social cognition or occupational mobility in a polycentric, multi-strafied industrial society. It would seem ill advised to clean up the experimental procedures by controlling for content or for any other stimulus variable. The price of such control is the creation of an unrealistic perceptual situation and a loss of generalizability beyond the laboratory. Content, voice quality, hesitations all impinge on the listener at once, and except in telephone conversations, so does the speaker's visual appearance. Add to these stimulus variables the listener variables of mood, personality, personal history, age and sex, and the lack of reliability comes as no surprise. Recent attempts to factor out variables in the stimulus to account for

differences in listener evaluation have not proven successful. Frederick Williams has concluded that listeners react to general speech and communication characteristics rather than specific linguistic cues.[10]

The movement called dialect modification did not arise from empirical research in social cognition, but instead from the intuitions of scholars trained in linguistics. If their scholarly intuitions about social cognition are not borne out by empirical research, the reason probably lies in the fact that, by virtue of their training, their perception of variations in speech is much more acute than that of the public at large. Along with teachers of speech and English, they seem to share a sensitization to the form of language which is largely a function of their professional activities.

The fact still remains that there is something different about the communication patterns of the poor. Recent linguistic debates notwithstanding, lower-class communication behavior is inadequate in the face of particular tasks. What nobody in the experiments heard was the tape left on the cutting-room floor. Among the lower class speakers, it was extremely difficult finding fifteen seconds of continuous discourse related to the interview question. Answers were either very brief or trailed off into irrelevancies until the interviewer prompted them back onto the subject. Another characteristic of these lower-class speakers was an extremely long lag between question and answer, sometimes ten seconds or more, with no signal to the interviewer that the question was being deliberated except for isolated "ahs" and fragmentary starts such as "well, I think..." These are deficiencies of communication behavior in the particular context of an interview with a stranger.

If we assume that communication behavior is the proper concern of public instruction, these findings suggest a redirection of special language instruction for the poor, and a redefinition of the problem. Dialect modification with its focus on mere phonology and syntax has defined the problem in terms far too narrow: what dialect modification attempts to change makes no consistent difference in social perception, not when dialect is mixed in with other speech variables. This is consistent with findings of both Victor Cline and Frederick Williams.

The broader context for instruction in communication is the speaking situation. The interview is such a speaking situation, and appropriate interview behaviors can be and have been taught.[11] Small group discussion among strangers on a general topic is again another speaking situation with a specific structure which is being taught. Presently, such communication skills are often a small part of the language arts curriculum, but the Speech Communication Association is currently working to assemble and amplify objectives for these communication behaviors into a single source. These objectives emphasize the want of speech communication as much as the how. The crippling limitation of dialect modification has been its narrow focus. The barrier postulate translated into a prediction about social cognition can probably be established

as true at a level of analysis broader than the sentence. That is where current linguistic analysis stops and, roughly, where communication analysis begins. Futher tests of the barrier postulate must always be translated into predictions about human social cognition, and must clear a few elementary statistical hurdles. The predictive inadequacy of the barrier postulate in its current narrow form suggests that dialect modification give way to communication training, for it is in the framework of the entire communication act that the real differences lie.

Footnotes

Mr. Lee is the Director of the American Language Institute at Florida State University. Part of this research was supported by OEG 4-9-190021-0032-057, from the U.S. Office of Education.

1. David P. Harris, "1970 TESOL Presidential Report to the Membership," TESOL Newsletter, IV (September/December 1970), p. 2

2. Raven I. McDavid, Jr., "A Checklist of Significant Features for Discriminating Social Dialects," in Eldonna L. Everttts, Dimensions of Dialect, National Council of Teachers of English (Champaign, 1967), pp. 7-10.

3. Victor B. Cline, "Person Perception from the Standpoint of an Empiricist," a paper read at the American Psychological Association Convention, San Francisco, September, 1968, p.6.

4. George W. Putnam and Edna M. O'Hearn, "The Status Significance of an Isolated Urban Dialect," Language Supplement, Language Dissertation No. 53, XXXI, (October-December, 1955).

5. William Labov, The Social Stratification of English in New York City, Center for Applied Linguistics, (Washington, D.C., 1966).

6. L.S. Harms, "Listener Judgments of Status Cues in Speech," Quarterly Journal of Speech, XLVII (1961), pp. 164-168.

7. Richard Tucker and Wallace E.Lambert, "White and Negro Listeners' Reactions to Various American-English Dialects," Social Forces, XLVII, (1969), pp. 463-468.

8. Robert L. Ebel, "Estimation of the Reliability of Ratings," Psychometrika, XVI (1951), pp. 407-424.

9. Martin Joos, "Language and the School Child, Word Study, VII (December, 1964). Cited in Virginia F. Allen, "Teaching Standard English as a Second Dialect," The Florida FL Reporter, VII (Spring/Summer, 1969) p. 123.

10. Frederick Williams, "Language, Attitude and Social Change," in Language and Poverty, ed. Frederick Williams, (Chicago, 1970), p. 388.

11. Richard R. Lee, "Linguistics, Communication and Behavioral Objectives: "Remedial Curriculum," Speech Teacher, XX (January, 1971). pp. 1-9.